Rec'd 12/16/68

SR 171

SENSITIVITY REDUCTION BY REOPTIMIZATION

Fred J. Taylor
Graduate Student
Department of Electrical Engineering
University of Colorado
Boulder, Colorado

THIS RESEARCH WAS SPONSORED BY THE NATIONAL SCIENCE FOUNDATION UNDER A N.S.F. TRAINEESHIP AND BY THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION, UNDER RESEARCH GRANT NGR 06-003-083

ACKNOWLEDGEMENTS

Thanks to Prof. I. M. Horowitz and Prof. H. Hermes for motivation, guidance and technical suggestions.

For 1969 JACC

Student Member IEEE

I. INTRODUCTION

Progress in control of the sensitivity of a system subject to parameter variations has not, in general, kept pace with advances achieved within the framework of optimal control. Considerable progress has been made in the design of linear systems with controlled sensitivity to parameter uncertainty. The foundation work in this class of problems is due to Bode [1], with applications and extensions to control systems by Horowitz [2,3]. Some methods have been presented in the field of optimal control systems which are subject to parameter variations over its interval of definition. A survey of these methods is given by M. Soboral Jr. [4]. Briefly, they will fall into the following two classes:

1. <u>Performance Sensitivity</u>

Consider any optimal control law which may be implemented either in an open loop or closed loop structure. What is the change in the cost index for the open and closed loop structures when parameter variations are considered [5,6]? Other variations of this idea exist in the form of min max values of the cost index over all parameter values. In Method 1 it is generally assumed that the parameters lie in a very small neighborhood of their nominal (fixed) values.

2. <u>Trajectory Sensitivity</u>

In this case there is generated a trajectory in the solution space, which, in the sense of an augmented given cost index, is least sensitive to parameter variations. The cost index is augmented by a term which is a measure of the trajectory dependence on parameter variations. This results, in general, in a non-optimal control for any specific set of parameter values with respect to the unaugmented cost index [7].

Kokototovic and Heller[1] have adopted an approach which preserves the concept of optimality. Their objective is a system which will be adaptive in the sense that the system will perform optimally for small parameter variations. However, they assume an apriori feedback structure, do not examine how small the 'small' parameter variations must be, and do not have a sensitivity analysis. They do not show that the adaptive structure can yield a cost which approaches arbitrary close to the true optimal cost. Also, their implementation restricts the number of parameters to the dimension of considered system, (although commenting that this computational problem may be overcome.)

This paper presents a method for designing a system which tries to operate optimally over its range of parameter variations. No structural constraints are apriori assumed; instead, a general mathematical formulation of the problem is presented, with the system's structural form being a derived result. Given a bounded set of allowable parameter variations, there is available a computable bound on the cost index variations.

<u>Symbols</u>

*	Optimal variables
e	Adaptive variables (sometimes referred to as "approximate optimal variables)
n	Nominal variables
′	Denotes transpose
$\langle , \rangle$	Non-degenerate inner product

[1]Received as a set of notes 1967.

II. Formulation of the Problem

Given the autonomous plant

$$\dot{x}(t) = f(x(t),u(t);\alpha) \quad x(t_o) = x_o, t\epsilon[t_o,T]$$

where $u(t)\ \epsilon\,\Omega \subset E^r$, $\Omega = \{u(t) = (u_1, \ldots, u_r)', \forall \text{ admissible } u_i\}$, $\alpha \triangleq \{(\alpha_1, \ldots, \alpha_m) \mid \alpha_i \epsilon \mathcal{A} \ \ i=1,\ldots, m\}\ \epsilon\ A \subset E^m$, $f : E^n \times A \times \Omega \to E^n$ continuously.

A performance index $C(u) = \int_{t_o}^{T} L\ (x_1 u)dt$ with $C(u)$: $E^n \times \Omega \to R$.

The optimization problem is to find a $u(t)$ which transfers the state of the plant from some initial state $x(t_o) = x_o$, to a point in a given target set S, such that C(u) is minimized.

Define: $H(x,p,u;\alpha) = L(x,u) + p\cdot f(x,u;\alpha)$ where $t\epsilon\ [t_o,T]$, $p\epsilon\ E^n$. (3)

In order that a control $u^*(t)$ be optimal it is necessary that $\exists$ a function $p^*(t) \ni$

(i) $p^*(t)$ and $u^*(t)$ are solutions of the canonical equations

$$\dot{x}^*(t) = \frac{\partial H}{\partial p}(x^*(t),\ p^*(t),\ x^*(t)\ ),\ x^*(t) = x_o \tag{4}$$

$$\dot{p}^*(t) = -\frac{\partial H}{\partial p}(x^*(t),\ p^*(t),u^*(t)) \tag{5}$$

(ii) $\min H[x^*(t),p^*(t),u(t)] = H[x^*(t),p^*(t),u^*(t)]\ u\epsilon\Omega$ (6)

(iii) and p(t) satisfies the usual transversality conditions dictated by the target set [9].

SET OF ADMISSIBLE PARAMETERS

Consider the continuity aspects of a system of first order O.D.E.'s for the following three cases:

For $x\epsilon E^n$

(i) $\dot{x}(t) = f(t,x(t)\ )$, $x(t_o) = y$, y considered to be a parameter,

(ii) $\dot{x}(t) = f(t,x,\alpha)$, $x(t_o) = x^o$, x^o fixed, $\alpha = (\alpha_1, \ldots, \alpha_m)'$ considered to be a parameter vector.

(iii) Combinations of (i) and (ii)

It is known [10] that under the proper reduction routine:

$$(i) <=> (ii) <=> (iii)$$

Moreover, the reductions preserve all the continuity properties between the classes. Because there is an equivalence between (i), (ii), and (iii), this paper will consider class (ii) formulations only.

For the sake of completeness, t_o may be considered a parameter in (i) if the latter's dimension is increased by one.

Proof: First convert the nonautonomous nth order O.D.E. to a (n+1) order autonomous O.D.E., by letting $\dot{x}_{n+1}(t) = 1$, $x_{n+1}(t_o) = t_o$, (by virtue of this transformation we shall restrict all future analysis to autonomous O.D.E.'s which may have originally been given as a nonautonomous O.D.E. of class (ii)). Let $z(\tau) = (x_{n+1}(\tau), x_1(\tau), \ldots, x_n(\tau))$ and $g(\tau,x) = (1, f(\tau,x))$. Then $\dot{z}(t) = g(z(\tau))$, $z(t_o) = (t_o, x_o)$. #

Define: $\mathcal{A}$ to be the set of all admissible parameters such that

$$\mathcal{A} = \{\alpha_i \mid |\alpha_i - \alpha_i^n| \le M_i, \; i = 1, \ldots, m, \text{ For some given } M_i\}$$

Define: A to be a parameter vector,

$$A = \{\alpha = (\alpha_1, \ldots, \alpha_m)' \mid \alpha_i \in \mathcal{A}, \; i = 1, \ldots, m\} \subset E^m.$$

If one considers a plant to be parameterized by a set of α's, $\alpha \in A$, then the optimization would have to be performed the cardinal number of A times. Therefore, one is motivated to seek an extension of an existing solution into a neighborhood of that solution. Consider an expansion of the Hamiltonian in a truncated Taylor Series about the nominal Hamiltonian. The latter corresponds to the optimal control $u^n(t)$ which satisfies (6) under the condition $\alpha = \alpha^n$. Because the Hamiltonian function is dependent upon the parameter vector which represents the existing parameters of the plant. The introduction of the general parameter vector α into the Hamiltonian shall be accomplished as follows:

Define a new state vector y, . $y \in E^n x A \subset E^{n+m}$ where $y(t) \triangleq \begin{bmatrix} x(t) \\ \alpha(t) \end{bmatrix}$, $x(t) \in E^n$, $\alpha \in A \subset E^m$, $y(t_o) = \begin{bmatrix} x_o \\ \alpha_o \end{bmatrix}$, where α_o is the initial parameter vector which, without any apriori knowledge of its value, will be assumed to be α^n.

Define the nominal Hamiltonian to be: $H^n(y(t), p(t), u(t)) = H(y^n(t), p^n(t), u^n(t))$, $p(t) \in E^{n+m}$

Theorem 1. Let H (y,p,u) and all its partial derivatives up through order k be continuous in a neighborhood N of (y^n, p^n, u^n). Then, for $(y, p, u) \in N$

$$H(y, p, u) = \sum_{i=0}^{k-1} \frac{1}{i!} \langle (y-y^n, p-p^n, u-u^n), \nabla \rangle^i H^n(y,p,u) +$$

$$\frac{1}{k!} \langle (y-y^n, p-p^n, u-u^n), \nabla \rangle^k H^q(y,p,u) \qquad (7)$$

where $H^q(y,p,u) = H(y^q, p^q, u^q)$, $(y^q, p^q, u^q) \in N$ and ∇ is the gradient operator [9]. (y^q, p^q, u^q) is a point on the line segment connecting $(y, p, u) \in N$ to $(y^n, p^n, u^n) \in N$. Geometrically: Let N be a subset of $E^{2(n+m)+r} \ni ||(y-y^n, p-p^n, u-u^n)|| \leq \epsilon \leq O$. $(y^q, p^q, u^q) \in co(N)$ (co denote convex hull). But N is compact therefore $N = co(N)$.

The last right hand term of (7) will be used to generate the truncated error which is later in the paper. For the moment, call the truncation error $o(\epsilon^k)$.

III. NECESSARY CONDITIONS FOR THE MINIMIZATION OF THE EXTENDED HAMILTONIAN SYSTEM

Consider the case where $k = 3$, $H(y,p,u) \epsilon C^4[t_o,T]$ in some neighborhood N. $C(u) = \int_{t_o}^{T} L(x,u)dt$ and the differential side constraint $\dot{y} = g\ (y,u)$, $y(t_o) = y_o$, given.

Notation: Let $\left[\frac{\partial^k H^n(y,p,u)}{\partial u^h \partial p^i \partial y^g}\right] = H^n{}_{y^j p^i u^k}$; $j + i + h = k$

Then from (4) and (7)

$$\dot{y} = \frac{\partial H(y,p,u)}{\partial p} = H_p^n\ (y,p,u) + H_{p^2}^n(y,p,u)(p-p^n) + H_{py}^n(y,p,u)(y-y^n)$$

$$+ H_{pu}^n(y,p,u)(u-u^n) + o_p(\epsilon^3) \tag{8}$$

From (5) and (7)

$$\dot{p} = -\frac{\partial H(y,p,u)}{\partial y} = -H_y^n\ (y,p,u) - H_{y^2}^n(y,p,u)(y-y^n) - H_{yp}^n(y,p,u)$$

$$(p-p^n) - H_{y^2}^n(y,p,u)(u-u^n) + o_y(\epsilon 3) \tag{9}$$

From (6) and (7), u satisfies

$$H_u(y,p,u) = H_u^n(y,p,u) + H_{uu}^n(y,p,u)(u-u^n) + H_{uy}^n(y,p,u)(y-y^n) +$$

$$H_{up}^n(y,p,u)(p-p^n) + o_u(\epsilon^3) \equiv o \text{ over } t\ \epsilon\ [t_o,T] \tag{10}$$

From equations (7), (8), (9), and (10) the following can be proven:

<u>Lemma 1.</u> The Taylor extension (7) satisfies $C(u) = C(u^n)$ if $\alpha = \alpha^n$ over $t\ \epsilon\ [t_o T]$.

Proof: Consider the sequence $\alpha_n \rightarrow \alpha^n$ uniformly, $\alpha_n \epsilon\ A$ which induces an isomorphic mapping σ: $y \rightarrow y_n$, $y_n = \begin{bmatrix} x \\ \alpha_n \end{bmatrix}$, $H(y,p,u) \epsilon\ C^3[t_o,T]$ given. Under σ, $\dot{y} = H_p(y,p,u)$, $y\ (t_o) = y_o$ becomes $\dot{y}_n = \ell(y_n)$ (11)

For p,u assumed known over $t \epsilon [t_o,T]$. Let ϖ^n be a solution of (11). $\varpi^n = y_o + \int_{t_o}^{t} \ell(\varpi_n)dt$, $\{\varpi^n\}$ is an equicontinuous family, bounded, which implies, by Ascoli's theorem, $\exists$ a subsequence $\{\varpi^n(k)\}$ converges uniformly, say to ϖ [10]. But $H_p(y,p,u)$ is Lipschitizian in y, $y \epsilon N$.

Therefore $\exists$ a unique solution to (11). Also, $\dot{p} = - H_y(\omega^{n(k)}, p, u)$,

$$p(T) = c \tag{12}$$

$H_y(\omega^{n(k)}, p, u)$ is Lipschitzian in p, $p \epsilon N$ then $\exists$ a unique solution to $\dot{p} = \bar{\ell}(p)$ for ω^n, u assumed given; call the unique solution $\bar{\omega}$.

For N sufficiently small, such that u^n is globally optimal for $\alpha = \alpha^n$, $t \epsilon [t_o, T]$, the implicit function theorem implies $\exists$ a unique

$$u \epsilon N \ni u = g\ (\omega^{n(k)}, \bar{\omega}) \tag{13}$$

As $n(k) \to \infty$ (ie: $\alpha_{n(k)} \to \alpha^n$) assume the unique solutions to (11), (12), and (13) are:

i) $\omega^{n(k)} = y^n$

ii) $\bar{\omega} = p^n$

iii) $u = u^n$ respectively. where (i), (ii) and (iii) uniquely satisfies a system of coupled equations. #

Lemma 2. The Taylor extension satisfies the inequality $C(u^*) \leq C(u^n)$ under the condition $\alpha = \alpha^* \neq \alpha^n$.

Proof: Suppose not, then $\exists$ a $u^n \epsilon N \ni C(u^n) < C(u^*)$. But for $(y^*, p^*, u^*) \epsilon N$, $\min_{u \epsilon \Omega} H(y^*, p^*, u) = H(y^*, p^*, u^*)$.

#

Define: $\Delta y = y - y^n$, $\Delta p = p - p^n$, $\Delta u = u - u^n$. Note also $H^n_u(y,p,u) = 0$ and $H^n_{pp}(y,p,u) = 0$. If $H^{n\,-1}_{uu}(y,p,u) \neq 0$, then, from 10, solving for Δu,

$$\Delta u = - H^{n\,-1}_{uu}(y,p,u)\left[H^n_{uy}(y,p,u)(y-y^n) + H^n_{up}(y,p,u)(p-p^n) + \omega_u(\epsilon^3)\right] \tag{14}$$

Substituting (10) into (8)

$$\begin{aligned} \Delta\dot{y} = {} & \left[H^n_{py}(y,p,u) - H^n_{pu}(y,p,u)\, H^{n\,-1}_{uu}(y,p,u)\, H^n_{uy}(y,p,u)\right](\Delta y) - \\ & \left[H^n_{pu}(y,p,u)\, H^{n\,-1}_{uu}(y,p,u)\, H^n_{up}(y,p,u)\right](\Delta p) - H^n_{pu}(y,p,u)\, H^{n\,-1}_{uu}(y,p,u) \\ & o_u(\epsilon^3) + o_p(\epsilon^3), \qquad \Delta y(t_o) = o \end{aligned} \tag{15}$$

Substituting (10) into (9)

$$\Delta\dot{p} = -\left[H^n_{yy}(y,p,u) - H^n_{yu}(y,p,u)\, H^{n\,-1}_{uu}(y,p,u)\, H^n_{uy}(y,p,u)\right](\Delta y)$$

$$+ \left[H^n_{yu}(y,p,u) - H^n_{yu}(y,p,u)\, H^n_{uu}(y,p,u)^{-1}\, H^n_{yp}(y,p,u)\right]\Delta p + H^n_{yu}$$

$$(y,p,u)\, H^n_{uu}(y,p,u)^{-1}\, O_u(\epsilon^3) - o_y(\epsilon^3), \quad \Delta p(T) = 0 \qquad (16)$$

$\exists\ N_\epsilon \subseteq N$ small enough $\ni O_u(\epsilon^3)$, $O_p(\epsilon^3)$, and $O_y(\epsilon^3)$ are negligible. Then (15) and (16) can be represented by a linear 2 (n+m) system of 1st order O.D.E's as follows

$$\begin{bmatrix} \Delta\dot{y} \\ \Delta\dot{p} \end{bmatrix} = \begin{bmatrix} A(t) & B(t) \\ C(t) & -A^T(t) \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta p \end{bmatrix} = G(t) \begin{bmatrix} \Delta y \\ \Delta p \end{bmatrix}, \quad \begin{bmatrix} \Delta y \\ \Delta p \end{bmatrix}_{t=t_o} = 0,$$ where the matrices $A(t)$, $B(t)$ and $C(t)$ are defined in the obvious way. (17)

Equation (17) is recognized to be a matrix Riccati differential equation. The solution is given by $K(t) = K'(t)$, $K(t)$ satisfies

$$-\dot{K}(t) = A'(t)\,K(t) + K(t)\,A(t) - K(t)\,B(t)K(t) - C(t), \quad K(T) = 0. \qquad (18)$$

where $p(t) = K(t)y(t)$ (19)

and upon substituting (19) into (14)

$$\Delta u = -H^n_{uu}(y,p,u)^{-1}\left[H^n_{uy}(y,p,u) + H^n_{up}(y,p,u)\,K\right]\Delta y \triangleq \tilde{\mathcal{G}}(t)\Delta y. [11] \qquad (20)$$

If any of the 2(n+m) solutions of (17) are known in closed form then reduction techniques may be used to reduce the computational problems in finding K(t). [8] The realization, in block diagram form, is given in Figure 1.

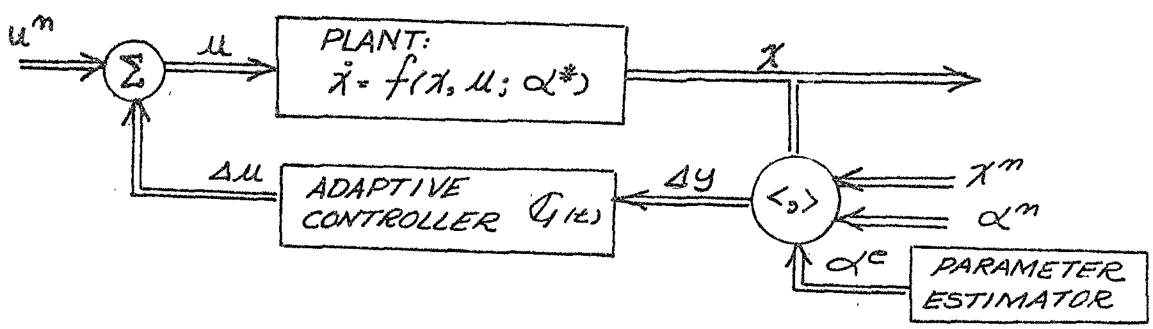

Figure 1

It is assumed (i) that the actual state $x(t)$ can be measured in real time, (ii) $x^n(t)$ and α^n can be loaded into the system in real time, and (iii) α can be generated. The question of parameter generation is discussed later in the paper. It may be noted that the feedback loop is adaptive and generates a control effort which is in such a direction as to minimize the cost incurred. This preservation of optimality will be called "reoptimization".

In order that the adaptive structure be of practical use it should possess the quality that $C(u^e) \le C(u^n)$ for a set of parameter vectors in some neighborhood on α^n.

Theorem 2: $\exists$ some neighborhood N_ϵ, $\ni$ for $(y,p,u) \epsilon N_\epsilon$, and $u^e \epsilon N_\epsilon \subset \Omega$, u^e satisfying (20) over $t \epsilon [t_o, T]$, then $C(u^*) \le C(u^n)$.

Proof: Part A; $C(u^*) \le C(u^e)$, let $\alpha = \alpha^* \ne \alpha^e$ apply Lemma 2. Part B; $C(u^e) \le C(u^n)$, suppose not, then $\exists\, N_\epsilon$ (ie: $\| y-y^n, p-p^n, u-u^n \| \le \epsilon) \ni$ $C(u^e) > C(u^n)$. But, $\exists\, \epsilon$, sufficiently small, say ϵ_1, $\epsilon_1 \ge 0 \ni H(y,p,u)$ $= H^n(y,p,u) + \sum_{i=1}^{2} \frac{1}{i!} \langle (y-y^n, p-p^n, u-u^n), \nabla \rangle^i H^n(y,p,u)$, (Implies truncation error approaches zero):

$$\text{Therefore } \epsilon_1 = \left\{ \begin{array}{ll} = 0 & \text{in which case equality holds in Part B. By Lemma 1 and } \epsilon_1 = 0 \text{ implies } \alpha=\alpha^n, \\ & \text{(ie: } H(y,p,u)\, \epsilon C^4[t_o,T] \text{ is not smooth), or} \\ > 0 & \text{implies } \alpha \ne \alpha^n \end{array} \right\}$$

If $\epsilon_1 = 0$ finished, if $\epsilon_1 > 0$, then for $(y,p,u) \epsilon N_{\epsilon_1}$ $\exists\, u^e = u^* \ni$ $H(y^e,p^e,u^e) = H(y^*,p^*,u^*)$, where $y^e(t)$ satisfies $(\dot{y}^e - \dot{y}^n) = A(t)$ $(y^e-y^n) + B(t)(p^e-p^n)$, $(y^e-y^n|_{t=t_o} = 0$ and $p^e(t)$ satisfies $(\dot{p}^e-\dot{p}^n) =$ $C(t)\ (y^e-y^n) + D(t)(p^e-p^n)$; $(p^e-p^n|_{t=T} = 0$. (ie: $y^*(t) = y^e(t)$ and $p^*(t) = p^e(t)$ #

IV. PARAMETER ESTIMATION

If the plant parameters can be monitored the "reoptimization" problem is direct. In general one cannot hope that all the parameters can be monitored. Koktovic and Heller[2] consider the following variational argument,

Let $\sigma_k^i \triangleq \partial x_i/\partial\alpha k$, $k = 1,\ldots,m$, and $\delta_x = \sigma\Delta\alpha$, if $m = n$ then $\Delta\alpha = \sigma^{-1}\delta x$. This dimensionality restriction can be removed by using a sufficient number of information bearing variables that are measurable in the system.[3] Suppose we can monitor directly the variables $x(t)$ and $u(t)$, $x(t)$ and $u(t)$ measurable functions, then we can estimate α as follows:

Consider $x(t)$ monitorable and satisfies $\dot{x}(t) = f(x,u;\alpha^*)$, $x(t_o) = x_o$, $u(t)$ monitorable and α not monitorable over $t \in [t_o,T]$. Assume its existence over $t \in [t - \ell\epsilon, t_1] \subseteq [t_o,T]$, $\ell,\epsilon > 0$, $t_1 > t_o$. (21)

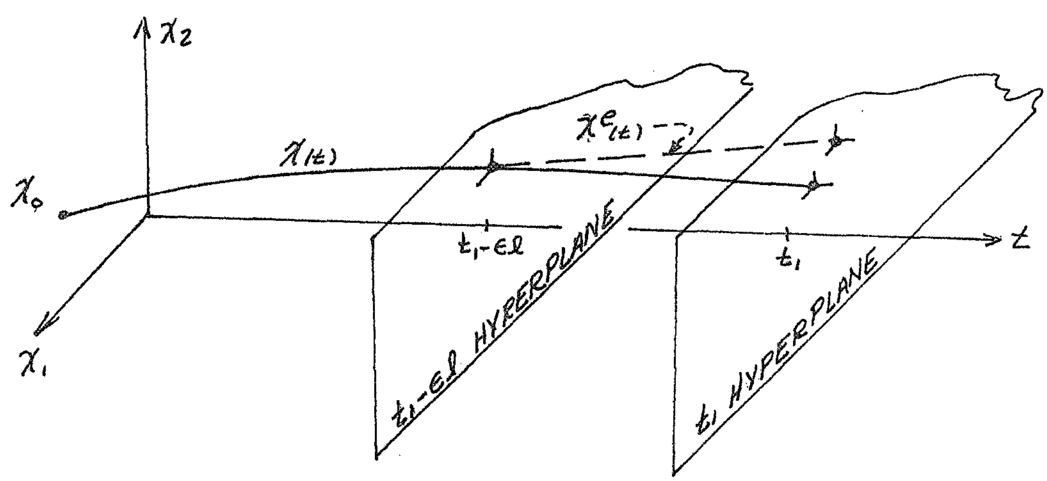

Fig. 2

[2] See Footnote 1.

[3] Conjecture: The posed estimation problem is: Equivalent to the estimation of stochastic parameters.

[4] Monitorable implies: can be obtained by direct measurement.

Consider a model of (21) to be $\dot{x}^e = f(x^e, u^e; \alpha^e)$, $x^e(t_1 - \ell\epsilon) = x(t_1 - \ell\epsilon)$ over $t \epsilon [t_1 - \ell\epsilon, t_1]$. (22)

For $x^e(t)$, $u^e(t)$ measurable, the game is to find an $\alpha^e \ni$ (22) models (21) in some sense of an E^n norm.[5] Define x_i^e to be the solution of (22) for $\alpha = \alpha_i^e$, $\alpha_i^e \epsilon A$, i = cardinality of A. Define $\Gamma_e(t) = \bigcup_{\forall \alpha_i^e \epsilon A} x_i^e(t)$, obviously $x(t_1) \cap \Gamma_e(t_1) \neq \phi$.

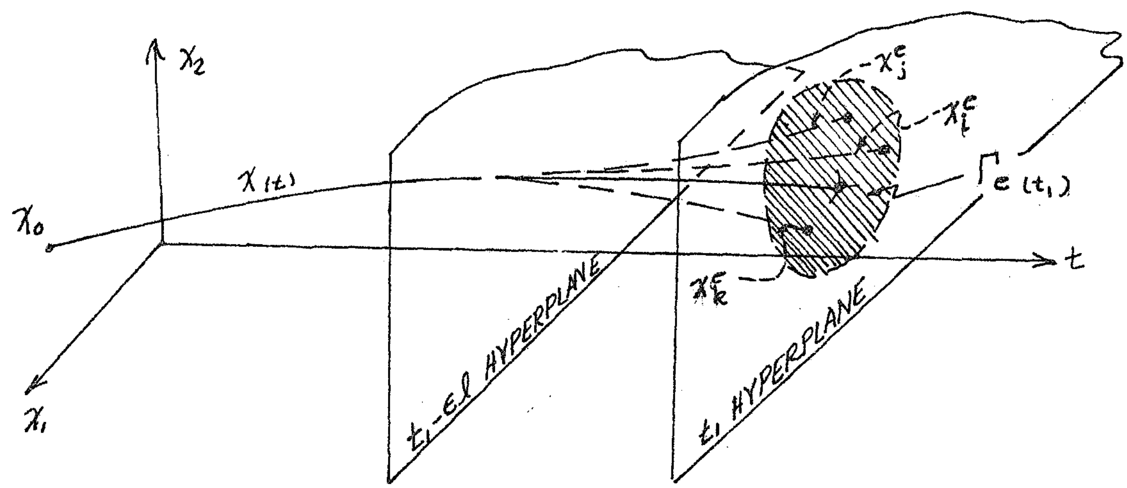

Fig. 3

Let us demand that t_1 is a Lebesque point of α, α measurable (ie: if $\int_{t_1-\ell\epsilon}^{t} f(\gamma, \sigma, \alpha^i)dt = f(\gamma(t_1), \sigma(t_1), \alpha^i)\ell\epsilon + o(\epsilon)$ (where $o(\epsilon)$ is: "order of ϵ"). Define an elementary perturbation of α as follows [12]:

$$\alpha_{\pi i} = \left\{ \alpha_i^e \text{ on } t_1 - \ell\epsilon \le t \le t_1; \alpha^* \text{ elsewhere} \right\} \ni \alpha^n \oplus \alpha_i^e = \alpha_{\pi_i} \quad (23)$$

Consider $\varphi_{(t)} \triangleq x^i(t,\epsilon)$, $u(t) \epsilon \Omega$ and assumed known by monitoring, and $x_i^e(t,\epsilon) = x(u, \alpha\pi_i)$. Now compute $\mathring{\varphi}(0)$ as follows: using (21), (22), (23), [13]

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon}[x_i^e(t_1,\epsilon) - x(t_1)] - \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{t_1-\ell\epsilon}^{t_1} (f(x_i^e(t,\epsilon), \alpha_i^e) - f(x(\tau),$$

[5] A $L^2(t_1 - \ell\epsilon, t_1)$ norm would be desirable. But the computational problems become severe.

$\alpha^*)\ )dt = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left\{ \left| f\ (x(t_1,\epsilon),\ \alpha_i^e) - f(x(t_1),\ \alpha^* \right| \ell\epsilon + O(\epsilon) \right\}$

$$= \left| f(x(t_1),\alpha_i^e) - f(x(t_1),\ \alpha^*) \right| \ell = \mathring{\omega}\ (0) \tag{24}$$

Define (i) $v\pi_i(t_1) = \frac{d}{d\epsilon}\ x\pi_i(t_1,\epsilon)\Big|_{\epsilon=0}$ (25)

(ii) $A\pi_i = \left\{ \alpha\pi_i \,|\, \alpha\pi_i \text{ close to } \alpha \right\}$ (26)

Then we can express $x\pi_i(t_1,\epsilon) = x(t_1) - v\pi_i(t_1)\epsilon + O(\epsilon)$ for $\alpha\pi_i \epsilon A_{\pi_i} \subset A$, where $x\pi_i(t_1,\epsilon)$ is a point function (27)

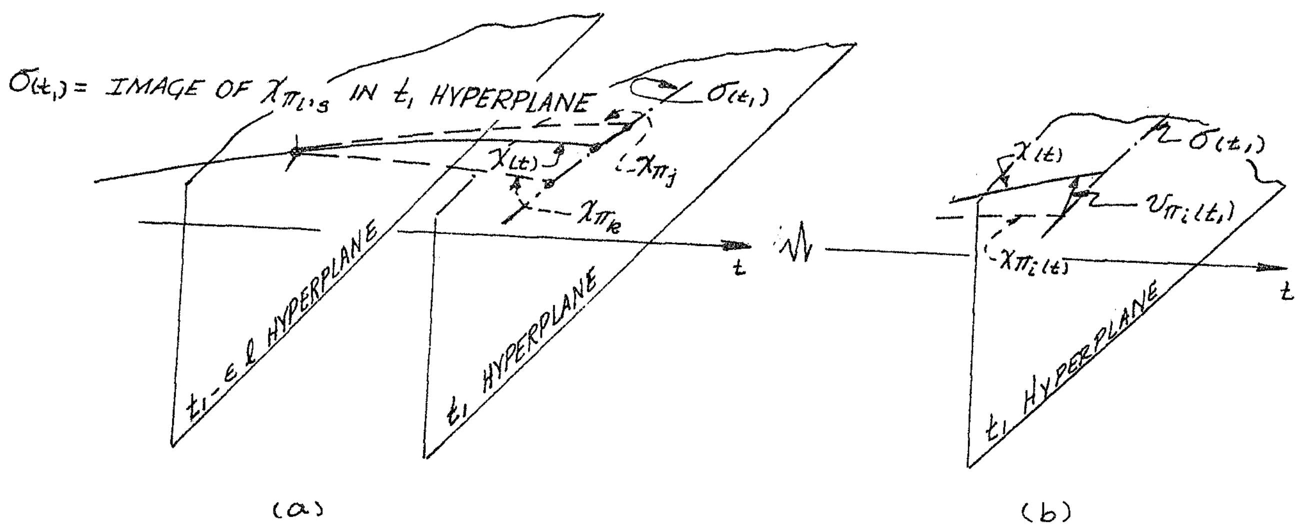

Fig. 4 (a,b)

$\sigma(t_1)$ is a hypersurface of all perturbed solutions for $\alpha\pi_i \epsilon A\pi_i$, $\nu\pi_i(t_1)$ is perpendicular to $\sigma(t_1)$ at the point $x\pi_i(t_1)$ if $O(\epsilon)$ is sufficiently small and $x\pi_i(t_1,\epsilon) \cong X(t_1) - V\pi_i(t_1)$ (28)

The measure of the error in approximating $x(t_1)$ with $x\pi_i(t_1,\epsilon)$ is $\|\ \nu\pi_i(t_1)\ \|$. Minimizing $\|\ \nu\pi_i(t_1)\|$ for ϵ fixed and sufficiently small. From (27) $(\nu\pi_i(t_1,\epsilon)) = (f(x(t),\alpha^*) - f(x\ t_1),\ \alpha_i^e)\ \epsilon =$ $(\dot{x}\ (t_1) - f(x(t_1),\alpha_i^e)\epsilon$ (29)

Now, several engineering problems arise from (29) and are: (i) Is $\dot{x}(t_1)$ monitorable? (ii) What are the dimensionality restriction? If $\dot{x}(t_1)$ is monitorable the α^e which minimizes (29) can be determined from a direct computation. For any, or all, of the $\dot{x}_i(t_1)$, $i = 1,\ldots,n$, which are not monitorable, $\dot{x}_i(t_1)$ must be calculated from available monitorable system variables. Some of the methods which will facilitate these are:

1) One-sided derivatives for x(t) sufficiently smooth and $\ell\epsilon$ sufficiently small, ℓ proportional to the apriori smoothness judgment on x(t).

2) If any element $x_i(t)$ of x(t) is not a function of a variable parameter then $\dot{x}_i(t)$ can be computed directly (ie: $\dot{x}_i(t) = g(x(t))$).

3) Mesh refinement methods where an observation interval is sampled at times t_i, $t_i \epsilon [t_j, t_k]$, $i = j, \ldots, k$, t_i monotonically increasing. Methods considered:

i) Differentiation of divided differences [14]

ii) Derivative formulas from difference operations [14]

iii) Central difference formulas [14]

iv) Modified Euler's Formula [14]

v) Weighted averaging of a sequence of $\dot{x}_i(t_{i's})$,($\dot{x}(t_i)$ found by any method) which will smooth the data and reduce the effect of data points which have a large variance from the mean.

Other methods which deserve attention are on-line gradient techniques, learning model techniques, and a method of considering the parameters to belong to a class of piecewise continuous functions, $\alpha\epsilon P[t_o,T]$, is currently being developed.[6]

[6] $\alpha\epsilon P[t_o,T]$ implies $\dot{\alpha} = o$ a.e. but the continuity everywhere of x(t) is preserved as noted by its integral equation representation.

Because the α_i^e's appear linearly in (29) we are interested in the solution to an equation of the form $A(t_1)\alpha_i^e = b(t_1)$ (30) where $b(t_1) = \dot{x}(t_1)$ and $A(t_1)$ is an nxm matrix of rank q. If $q = m = n$, the solution of (30) is $\alpha_i^e = A^{-1}(t_1)b(t_1)$. Otherwise, two general kinds of degeneracy can occur. R. E. Mortensen [15] listed them as, 1; If $b(t_1) \epsilon R(A(t_1))$, then no exact solution is possible. 2; if $N(A(t_1)) \neq \phi$, then a solution, if it exists, is not unique.[7] Case (1) might arise if the methods used to compute $b(t_1)$ are considered to be noisy. Since an exact solution is not possible we may request a "best least squares fit", Consider a new set of equations. $\tilde{A}(t_1)\alpha_i^e = \tilde{C}(t_1)$ where $\tilde{C}(t_1) = \{[\cdots \dot{x}_i(t_1)\cdots]'$ $1 \leq i \leq n$ and $\Sigma i = m\}$, where $\tilde{A}(t_1)$ is mxm rank q, $m < n$, $\alpha_i^e \epsilon A \subset E^m$, $\tilde{C}(t_1) \epsilon E^n$, and if $q = m$ then a solution exists. If there are more unknowns than equations, then consider the following quadratic programming problem with an ordered vector cost functional will be considered:

Write $A(t_1)\alpha_i^e - b(t_1) = e$, $\alpha \epsilon A \subset E^m; b(t_1), e \epsilon E^m$. $||\alpha_i^e - \alpha^n||^2 = \langle\alpha_i^e - \alpha^n\rangle' \langle\alpha_i^e - \alpha^n\rangle$ and $||e||^2 = e'e$. Put $\Gamma = \min\{||e|| \succ ||\alpha_i^e - \alpha^n||\}$. This means first determine $\alpha_i^e \ni ||e||$ is minimized.

If $b(t_1) \epsilon R(A(t_1))$, then $||e|| = 0$, otherwise min. $||e|| > 0$. If $N(A(t_1)) = \phi$ then the solution is unique and problem finished. If $N(A(t_1)) \neq \phi$ then it is possible to minimize not only $||e||$ but $||\alpha_i^e - \alpha^n||$ also, which results in a unique solution α_i^e.

<u>Theorem 3:</u> The solution to $\min\{||e|| \succ ||\alpha_i^e - \alpha^n||\}$, where $A(t_1)\alpha_i^e - b(t_1) = e$, ($A(t_1)$, $b(t_1)$, α^n given) is $\alpha_i^e = A^\dagger(t_1)b(t_1)$, where $A^\dagger$ is the pseudioinverse.

[7] (1) implies more equations than unknowns. (2) implies more unknowns than equations.

Another technique which can be utilized for the problem of more unknowns than equations, if α_i^e is assumed to be essentially constant over $t_p \epsilon [t_i, t_k] \subset [t_o, T]$, is to define for $\alpha_i^e \epsilon A \subset E^m$, $b(t_p) \epsilon E^n$, $m > n$, and form the equation $\hat{A}(t_m)\alpha_i^e = \hat{b}(t_m)$ as follows:

$$\begin{bmatrix} A(t_\ell)\alpha_i^e \\ A(t_{\ell+1})\alpha_i^e \\ \vdots \\ A(t_{\ell+v-1})\alpha_i^e \\ \bar{A}(t_{\ell+v})\alpha_i^e \end{bmatrix} = \begin{bmatrix} b(t_\ell) \\ b(t_{\ell+1}) \\ \\ b(t_{\ell+v-1}) \\ \bar{b}(t_{\ell+v}) \end{bmatrix} ; \quad \begin{matrix} t_\ell \epsilon [t_i, t_k], b(t_\ell) \epsilon E^n \\ t_{\ell+1} \epsilon [t_i, t_k],\ b(t_{\ell+1}) \epsilon E^n \\ \\ t_{\ell+v-1} \epsilon [t_i, t_k],\ b(t_{\ell+v-1}) \epsilon E^n \\ t_{\ell+v} \epsilon [t_i, t_k],\ b(t_{\ell+v}) \epsilon E^w \end{matrix}$$

$\ni$ Rank $[\hat{A}(t_m)] = m$, then there exists a unique solution $\alpha_i^e = \hat{A}(t_m)^{-1}\hat{b}(t_m)$

V. FURTHER TOPICS

(1). Error Analysis: It was previously noted that for $(y-y^n, p-p^n, u-u^n) \in N$, then the truncation error of (7) is $\frac{1}{3!}$ $< (y-y^n),(p-p^n),(u-u^n), \nabla >^3 H^Q(y,p,u)$ for $(y^q-y^n, p^q-p^n, u^q-u^n) \in N$. (31) The truncation error is bounded by $\sup_{t\in[t_0,T]} \{(31)\}$. For a sufficiently small neighborhood N the truncation error can be made as small as desired. If $\dot{x}(t) = A(\alpha)x(t) + bu(t), x(0) = x_0$ and $C(u) = \frac{1}{2}\int_{t_0}^{T} (\langle x,Qx\rangle + \langle u,Ru\rangle)\, dt$; $Q \geq 0$, $R > 0$, $u(t)\in E^2, x(t)\in E^2$, $[A(\alpha)] = \begin{bmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{bmatrix}$ and $b = \begin{bmatrix}0\\1\end{bmatrix}$, then if we choose $n \ni (y-y^n, p-p^n, u-u^n) \leq \epsilon$ it can be shown that the maximum truncation error is less than $\frac{11}{3}\epsilon^3$. Also it can be shown by theorem 2, (19), and that for a sufficiently small N, that as $\alpha_n^e \to \alpha^*$, as $n \to \infty$, for a sufficiently large $n, \exists\ ||p_n^e - p^n|| \leq ||K(t)||\ ||y_n^e - y^n||$ and $||u_n^e - u^n|| \leq ||G(t)||$

$||y_n^e - y^n|| + ||H(t)||\ ||p_n^e - p^n|| \leq [||G(t)|| + ||H(t)||\ ||K(t)||]$

$||y_n^e - y^n||$ where G(t) and H(t) are defined by (14). Therefore, if $\exists\ N \ni ||y-y^n, p-p^n, u-u^n|| \leq \epsilon \ni$ the truncation error is sufficiently small, then we can choose ϵ to satisfy: $(1 + ||K(t)||\ (1 + ||H(t)||) + ||G(t)||)\ ||y-y^n|| \triangleq ||J(t)||\ ||y-y^n|| \leq \epsilon$. If $||\alpha - \alpha^n||$ is chosen to be less than some ϵ_1, $\epsilon_1 \geq 0$, then $||x - x^n|| \leq \epsilon\ ||J(t)||^{-1} - \epsilon_1$.

(2). Local Sensitivity: For the conditions given in Theorem 2, we can be assured of an improvement in the system sensitivity if the sensitivity measure $S(u^k)$ is defined to be $S(u^k) = C(u^k) - C(u^*) \geq 0$. (31) Under these conditions $S(u^e) \leq S(u^n)$, or the variations from the true optimal cost of the adaptive system is less than, or equal to, that of the open loop system using $u = u^n$ as a control effort.

VI. SUMMARY

This paper has established an adaptive system which was based on the solution of a truncated Hamiltonian system of equations. The only restrictions placed on the problem's formulation was that the Hamiltonian belonged to class $C^3[t_o,T]$ and $H^n_{uu}(y,p,u)^{-1} \neq 0$. From the approximate Hamiltonian, (ie: including bilinear terms), the adaptive control effort was found to be linear. The adaptive control was shown to, under certain neighborhood restrictions, reoptimize the system. A reduction in the system's sensitivity to parameter variations, over a system using nominally optimal control only, was gained by reoptimization.

BIBLIOGRAPHY

[1] Bellman, R. and Kalaba, R. ed., Selected Papers on Mathematical Trends in Control Theory, Dover, pp. 106-123, 1964.

[2] Horowitz, I. M., "Fundamental Theory on Automatic Linear Feedback Control Systems," IRE Transactions on Automatic Control, Vol. [illegible], AC-4, pp. 5-19, Dec. 1959.

[3] Horowitz, I. M., Synthesis of Feedback Systems, Academic Press, 1963.

[4] Sorbal, M., Jr., "Sensitivity in Optimal Control Systems", Proceedings of the I.E.E.E., Vol. 56, No. 10, pp. 1644-1652, Oct. 1968.

[5] Doroto, P., "On the Sensitivity in Optimal Control Systems," I.E.E.E. Transactions on Automatic Control, Vol. AC-8, No. 3, pp. 256-257, July 1963.

[6] Pagurek, B., "Sensitivity of the Performance of Optimal Control Systems to Parameter Variations," I.E.E.E. Transactions on Automatic Control, Vol. AC-10, pp. 178-180, April 1965.

[7] Bradt, H., "The Design of Optimal Controllers to Minimize A Performance Index Containing Sensitivity Functions," Ph.D. Dissertation, Dept. of Elect. Eng., Univ. of Denver, Oct. 1967.

[8] Hartman, P., Ordinary Differential Equations, John Wiley & Sons, pp. 49, 93, 94, 1964.

[9] Fulks, W., Advanced Calculus, Wiley, p. 230, 1961.

[10] Simmons, G., Introduction to Topology and Modern Analysis, McGraw-Hill, p. 126, 1963.

[11] Athans and Falb, Optimal Control, McGraw-Hill, pp. 761-766, 1966.

[12] Pontryagin, Boltyanski, Gamkrelidze, Mishchenko, Mathematical Theory of Optimal Control Processes, Interscience Pub., p. 75, 1962.

[13] Livsternik and Sobolev, Elements of Functional Analysis, Ungar, pp. 182-187, 1961.

[14] Kelly, Handbook of Numerical Methods and Applications, Addison Wesley, Ch. 4 and Ch. 7, 1967.

[15] Mortensen, R. E., "A Note On Polar Decomposition and the Generalized Inverses of Arbitrary Matrix," Set Of Notes Under Grant No. AF-AFOSR-139-64.